

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-49383

(P2002-49383A)

(43) 公開日 平成14年2月15日 (2002.2.15)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テマコード (参考)
G 1 0 L 13/00		H 0 3 M 7/30	Z 5 J 0 6 4
H 0 3 M 7/30		G 1 0 L 9/14	M
		9/12	M

審査請求 未請求 請求項の数16 O L (全 10 頁)

(21) 出願番号 特願2000-238892 (P2000-238892)

(22) 出願日 平成12年8月2日 (2000.8.2)

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 近藤 哲二郎

東京都品川区北品川6丁目7番35号ソニー株式会社内

(72) 発明者 渡辺 勉

東京都品川区北品川6丁目7番35号ソニー株式会社内

(74) 代理人 100082740

弁理士 田辺 恵基

最終頁に続く

(54) 【発明の名称】 デジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体

(57) 【要約】

【課題】 デジタル信号の波形再現性を一段と向上し得るデジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体を提案する。

【解決手段】 デジタルオーディオ信号D10の極性に基づいてデジタルオーディオ信号D10のクラスを分類し、当該分類されたクラスに対応した予測方式でデジタルオーディオ信号D10を変換するようにしたことにより、一段とデジタルオーディオ信号D10の特徴に適応した変換を行うことができる。

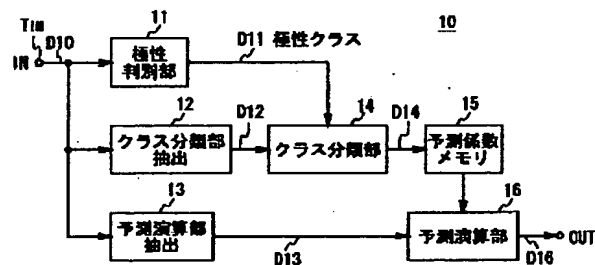


図1 オーディオ信号処理装置の構成

## 【特許請求の範囲】

【請求項1】 デジタルオーディオ信号を変換するデジタル信号処理装置において、

上記デジタルオーディオ信号の極性を判別する極性判別手段と、

上記極性判別結果に基づいて上記デジタルオーディオ信号のクラスを分類するクラス分類手段と、

上記分類されたクラスに対応した予測方式で上記デジタルオーディオ信号を予測演算することにより上記デジタルオーディオ信号を変換してなる新たなデジタルオーディオ信号を生成する予測演算手段とを具えることを特徴とするデジタル信号処理装置。

【請求項2】 上記極性判別手段は、上記デジタルオーディオ信号を時間軸領域に分割し、各分割領域ごとにその極性を判別することを特徴とする請求項1に記載のデジタル信号処理装置。

【請求項3】 上記極性判別手段は、上記デジタルオーディオ信号を少なくとも、正領域のみ、負領域のみ及びゼロクロスを含む領域の3つの領域に極性クラス分けすることを特徴とする請求項1に記載のデジタル信号処理装置。

【請求項4】 上記予測演算手段は、予め所望とするデジタルオーディオ信号に基づいて学習により生成されている予測係数を用いることを特徴とする請求項1に記載のデジタル信号処理装置。

【請求項5】 デジタルオーディオ信号を変換するデジタル信号処理方法において、

上記デジタルオーディオ信号の極性を判別する極性判別ステップと、

上記極性判別結果に基づいて上記デジタルオーディオ信号のクラスを分類するクラス分類ステップと、

上記分類されたクラスに対応した予測方式で上記デジタルオーディオ信号を予測演算することにより上記デジタルオーディオ信号を変換してなる新たなデジタルオーディオ信号を生成する予測演算ステップとを具えることを特徴とするデジタル信号処理方法。

【請求項6】 上記極性判別ステップでは、上記デジタルオーディオ信号は時間軸領域に分割され、各分割領域ごとにその極性が判別されることを特徴とする請求項5に記載のデジタル信号処理方法。

【請求項7】 上記極性判別ステップでは、上記デジタルオーディオ信号は少なくとも、正領域のみ、負領域のみ及びゼロクロスを含む領域の3つの領域に極性クラス分けされることを特徴とする請求項5に記載のデジタル信号処理方法。

【請求項8】 上記予測演算ステップでは、予め所望とするデジタルオーディオ信号に基づいて学習により生成されている予測係数が用いられることを特徴とする請求項5に記載のデジタル信号処理方法。

【請求項9】 デジタルオーディオ信号を変換するデ

ジタル信号処理装置の上記変換処理の予測演算に用いられる予測係数を生成する学習装置において、

所望とするデジタルオーディオ信号から当該デジタルオーディオ信号を劣化させた生徒デジタルオーディオ信号を生成する生徒デジタルオーディオ信号生成手段と、

上記生徒デジタルオーディオ信号の極性を判別する極性判別手段と、

上記判別された極性に基づいて上記生徒デジタルオーディオ信号のクラスを分類するクラス分類手段と、

上記デジタルオーディオ信号と上記生徒デジタルオーディオ信号とに基づいて上記クラスに対応する予測係数を算出する予測係数算出手段とを具えることを特徴とする学習装置。

【請求項10】 上記極性判別手段は、上記デジタルオーディオ信号を時間軸領域に分割し、各分割領域ごとにその極性を判別することを特徴とする請求項9に記載の学習装置。

【請求項11】 上記極性判別手段は、上記デジタルオーディオ信号を少なくとも、正領域のみ、負領域のみ及びゼロクロスを含む領域の3つの領域に極性クラス分けすることを特徴とする請求項9に記載の学習装置。

【請求項12】 デジタルオーディオ信号を変換するデジタル信号処理装置の上記変換処理の予測演算に用いられる予測係数を生成する学習方法において、

所望とするデジタルオーディオ信号から当該デジタルオーディオ信号を劣化させた生徒デジタルオーディオ信号を生成する生徒デジタルオーディオ信号生成ステップと、

上記生徒デジタルオーディオ信号の極性を判別する極性判別ステップと、

上記判別された極性に基づいて上記生徒デジタルオーディオ信号のクラスを分類するクラス分類ステップと、

上記デジタルオーディオ信号と上記生徒デジタルオーディオ信号とに基づいて上記クラスに対応する予測係数を算出する予測係数算出ステップとを具えることを特徴とする学習方法。

【請求項13】 上記極性判別ステップでは、上記デジタルオーディオ信号は時間軸領域に分割され、各分割領域ごとにその極性が判別されることを特徴とする請求項12に記載の学習方法。

【請求項14】 上記極性判別ステップでは、上記デジタルオーディオ信号は少なくとも、正領域のみ、負領域のみ及びゼロクロスを含む領域の3つの領域に極性クラス分けされることを特徴とする請求項12に記載の学習方法。

【請求項15】 デジタルオーディオ信号の極性を判別する極性判別ステップと、

上記極性判別結果に基づいて上記デジタルオーディオ信号のクラスを分類するクラス分類ステップと、

10

20

30

40

50

上記分類されたクラスに対応した予測係数を用いて上記デジタルオーディオ信号を予測演算することにより上記デジタルオーディオ信号を変換してなる新たなデジタルオーディオ信号を生成する予測演算ステップとを含むプログラムをデジタル信号処理装置に実行させるプログラム格納媒体。

【請求項16】 所望とするデジタルオーディオ信号から当該デジタルオーディオ信号を劣化させた生徒デジタルオーディオ信号を生成する生徒デジタルオーディオ信号生成ステップと、

上記生徒デジタルオーディオ信号の極性を判別する極性判別ステップと、

上記判別された極性に基づいて上記生徒デジタルオーディオ信号のクラスを分類するクラス分類ステップと、上記デジタルオーディオ信号と上記生徒デジタルオーディオ信号とに基づいて上記クラスに対応する予測係数を算出する予測係数算出ステップとを含むプログラムを学習装置に実行させるプログラム格納媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明はデジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体に関し、レートコンバータ又はPCM(Pulse Code Modulation) 復号装置等においてデジタル信号に対してデータの補間処理を行うデジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体に適用して好適なものである。

【0002】

【従来の技術】 従来、デジタルオーディオ信号をデジタル/アナログコンバータに入力する前に、サンプリング周波数を元の値の数倍に変換するオーバーサンプリング処理を行っている。これにより、デジタル/アナログコンバータから出力されたデジタルオーディオ信号はアナログ・アンチ・エイリアス・フィルタの位相特性が可聴周波数高域で一定に保たれ、また、サンプリングに伴うデジタル系のイメージ雑音の影響が排除されるようになされている。

【0003】 かかるオーバーサンプリング処理では、通常、線形一次(直線)補間方式のデジタルフィルタが用いられている。このようなデジタルフィルタは、サンプリングレートが変わったりデータが欠落した場合等に、複数の既存データの平均値を求めて直線的な補間データを生成するものである。

【0004】

【発明が解決しようとする課題】 ところが、オーバーサンプリング処理後のデジタルオーディオ信号は、線形一次補間によって時間軸方向に対してデータ量が数倍に緻密になっているものの、オーバーサンプリング処理後のデジタルオーディオ信号の周波数帯域は変換前とあまり変わらず、音質そのものは向上していない。さらに、補

間されたデータは必ずしもA/D変換前のアナログオーディオ信号の波形に基づいて生成されたのではないため、波形再現性もほとんど向上していない。

【0005】 また、サンプリング周波数の異なるデジタルオーディオ信号をダビングする場合において、サンプリング・レート・コンバータを用いて周波数を変換しているが、かかる場合でも線形一次デジタルフィルタによって直線的なデータの補間しか行うことができず、音質や波形再現性を向上することが困難であった。さらに、デジタルオーディオ信号のデータサンプルが欠落した場合において同様である。

【0006】 本発明は以上の点を考慮してなされたもので、デジタル信号の波形再現性を一段と向上し得るデジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体を提案しようとするものである。

【0007】

【課題を解決するための手段】 かかる課題を解決するため本発明においては、デジタルオーディオ信号の極性に基づいてデジタルオーディオ信号のクラスを分類し、当該分類されたクラスに対応した予測方式でデジタルオーディオ信号を変換するようにしたことにより、一段とデジタルオーディオ信号の特徴に適応した変換を行うことができる。

【0008】

【発明の実施の形態】 以下図面について、本発明の一実施の形態を詳述する。

【0009】 図1においてオーディオ信号処理装置10は、デジタルオーディオ信号(以下これをオーディオデータと呼ぶ)のサンプリングレートを上げたり、オーディオデータを補間する際に、真値に近いオーディオデータをクラス分類適用処理によって生成するようになされている。因みに、デジタルオーディオ信号とは、人や動物が発する声を表す音声信号、楽器が発する楽音を表す楽音信号、及びその他の音を表す信号を意味するものである。

【0010】 すなわち、オーディオ信号処理装置10において、極性判別部11は入力端子 $T_{in}$ から供給された図2に示す入力オーディオデータ $D_{10}$ を所定時間毎の領域(この実施の形態の場合、例えば6サンプル毎とする)に分割した後、当該分割された各時間領域の波形について、図2に示す極性判別方法によりその極性クラスを判別する。

【0011】 すなわち図2において、分割された領域AR1(カレントデータC1)のように切り出されたすべてのタップが正である場合、この極性クラスをCLASS0とし、分割された領域AR2(カレントデータC2)のように切り出された領域内にゼロクロスが存在すると共にカレントデータ(C2)が正である場合、この極性クラスをCLASS1とし、分割された領域AR4(カレントデータC4)のように切り出された切り出された領域内

にゼロクロスが存在すると共にカレントデータ(C4)が負である場合、この極性クラスをCLASS2とし、分割された領域AR3(カレントデータC3)のように切り出されたすべてのタップが負である場合、この極性クラスをCLASS3とする。

【0012】このように、オーディオデータD10の極性に基づく極性クラスを設定することにより、オーディオデータD10がゼロレベル近傍である場合の音素と大振幅部での音素とを正負両方の領域で一段と明確に区別することができる。

【0013】極性判別部11は入力オーディオデータD10のこのときのカレントデータに対応して求められた極性判別結果(CLASS0、CLASS1、CLASS2又はCLASS3)を極性クラスデータD11としてクラス分類部14に供給する。

【0014】また、クラス分類部抽出部12は入力端子T<sub>in</sub>から供給された入力オーディオデータD10を、極性判別部11の場合と同様の時間領域(この実施の形態の場合例えば6サンプル)に分割することによりクラス分類しようとするオーディオ波形データD12を抽出し、これをクラス分類部14に供給する。

【0015】クラス分類部14は、クラス分類抽出部12において切り出されたオーディオ波形データD12について、当該オーディオ波形データD12を圧縮して圧縮データパターンを生成するADRC(Adaptive Dynamic Range Coding)回路部と、オーディオ波形データD1 \*  
DR=MAX-MIN+1

$$Q = \{ (L - MIN + 0.5) \times 2^* / DR \} \dots\dots (1)$$

【0020】に従って、領域内の最大値MAXと最小値MINとの間を指定されたビット長で均等に分割して量子化を行う。なお、(1)式において{ }は小数点以下の切り捨て処理を意味する。かくしてオーディオ波形上の6つの波形データが、それぞれ例えば8ビット(m=8)で構成されているとすると、これらはADRC回路部においてそれぞれが2ビットに圧縮される。 ※

$$class = \sum_{i=1}^n q_i (2^P)^i$$

【0023】に示す演算を実行することにより、そのブロック(q<sub>1</sub> ~ q<sub>6</sub>)が属するクラスを示すクラスコードclassを算出すると共に、当該算出されたオーディオ波形データD12に基づくクラスコードclassに対して、上述の極性クラスCLASSを統合した後、当該統合されたクラスコードclassを表すクラスコードデータD14を予測係数メモリ15に供給する。このクラスコードclassは、予測係数メモリ15から予測係数を読み出す際の読み出しアドレスを示す。因みに(2)式において、nは圧縮されたオーディオ波形データq<sub>i</sub>の数を表し、この実施の形態の場合n=6であり、またPはビット割り当てを表し、この実施の形態の場合P=2であ

\*2の属するクラスコードを発生するクラスコード発生回路部とを有する。

【0016】ADRC回路部はオーディオ波形データD12に対して、例えば8ビットから2ビットに圧縮するような演算を行うことによりパターン圧縮データを形成する。このADRC回路部は、適応的量子化を行うものであり、ここでは、信号レベルの局所的なパターンを短い語長で効率的に表現することができるので、信号パターンのクラス分類のコード発生用に用いられる。

10 【0017】具体的には、オーディオ波形上の6つの8ビットのデータ(オーディオ波形データ)をクラス分類しようとする場合、2<sup>18</sup>という膨大な数のクラスに分類しなければならず、回路上の負担が多くなる。そこで、この実施の形態のクラス分類部14ではその内部に設けられたADRC回路部で生成されるパターン圧縮データに基づいてクラス分類を行う。例えば6つのオーディオ波形データに対して1ビットの量子化を実行すると、6つのオーディオ波形データを6ビットで表すことができ、2<sup>6</sup>=64クラスに分類することができる。

20 【0018】ここで、ADRC回路部は、切り出された領域内のオーディオ波形のダイナミックレンジをDR、ビット割り当てをm、各オーディオ波形データのデータレベルをL、量子化コードをQとすると、次式、

【0019】

【数1】

30※【0021】このようにしてダイナミックレンジで正規化され圧縮されたオーディオ波形データをそれぞれq<sub>i</sub>(n=1~6)とすると、クラス分類部14に設けられたクラスコード発生回路部は、圧縮されたオーディオ波形データq<sub>i</sub>に基づいて、次式、

【0022】

【数2】

$$\dots\dots (2)$$

る。

40 【0024】このようにして、クラス分類部14はクラス分類部抽出部12において入力オーディオデータD10から切り出されたオーディオ波形データD12そのもののクラスコードclassと、オーディオ波形データD12の極性クラスCLASSとを統合したクラスコードデータ(class')D14を生成し、これを予測係数メモリ15に供給する。因みに、オーディオ波形データD12そのもののクラスコードclassと、オーディオ波形データD12の極性クラスCLASSとを統合する方法として、クラス分類部14は例えばオーディオ波形データD12そのもののクラスコードclassに極性クラスCLASSを付加

することにより、これらを統合することができる。

【0025】予測係数メモリ15には、各クラスコードに対応する予測係数のセットがクラスコードに対応するアドレスにそれぞれ記憶されており、クラス分類部14から供給されるクラスコードデータD14に基づいて、当該クラスコードに対応するアドレスに記憶されている予測係数のセット $w_1 \sim w_n$ が読み出され、予測演算部16に供給される。

$$y' = w_1 x_1 + w_2 x_2 + \dots + w_n x_n \quad \dots (3)$$

【0028】に示す積和演算を行うことにより、予測結果 $y'$ を得る。この予測値 $y'$ が、音質が改善されたオーディオデータD16として予測演算部16から出力される。

【0029】なお、オーディオ信号処理装置10の構成として図1について上述した機能ブロックを示したが、この機能ブロックを構成する具体的構成として、この実施の形態においては図3に示すコンピュータ構成の装置を用いる。すなわち、図3において、オーディオ信号処理装置10は、バスBUSを介してCPU21、ROM(Read Only Memory)22、予測係数メモリ15を構成するRAM(Random Access Memory)15、及び各回路部がそれぞれ接続された構成を有し、CPU11はROM22に格納されている種々のプログラムを実行することにより、図1について上述した各機能ブロック(極性判別部11、クラス分類部抽出部12、予測演算部抽出部13、クラス分類部14及び予測演算部16)として動作するようになされている。

【0030】また、オーディオ信号処理装置10にはネットワークとの間で通信を行う通信インターフェース24、フロッピディスクや光磁気ディスク等の外部記憶媒体から情報を読み出すリムーバブルドライブ28を有し、ネットワーク経由又は外部記憶媒体から図1について上述したクラス分類適用処理を行うための各プログラムをハードディスク装置25のハードディスクに読み込み、当該読み込まれたプログラムに従ってクラス分類適用処理を行うこともできる。

【0031】ユーザは、キーボードやマウス等の入力手段26を介して種々のコマンドを入力することにより、CPU21に対して図1について上述したクラス分類処理を実行させる。この場合、オーディオ信号処理装置10はデータ入出力部27を介して音質を向上させようとするオーディオデータ(入力オーディオデータ)D10を入力し、当該入力オーディオデータD10に対してクラス分類適用処理を施した後、音質が向上したオーディオデータD16をデータ入出力部27を介して外部に出力し得るようになっている。

【0032】因みに、図4はオーディオ信号処理装置10におけるクラス分類適用処理の処理手順を示し、オーディオ信号処理装置10はステップSP11から当該処理手順に入ると、続くステップSP12において入力オ

\*【0026】予測演算部16は、予測演算部抽出部13において入力オーディオデータD10から時間軸領域で切り出された予測演算しようとするオーディオ波形データ(予測タップ)D13( $x_1 \sim x_n$ )と、予測係数 $w_1 \sim w_n$ に対して、次式

【0027】

【数3】

オーディオデータD10の極性を極性判別部11において算出する。

【0033】この算出された極性はオーディオ波形データD12のクラス分類を一段と確実にするためのものであり、オーディオ信号処理装置10は、ステップSP13においてクラス分類部14によりオーディオ波形データD12及び極性クラスD11に基づいてオーディオ波形データD12をクラス分類する。そしてオーディオ信号処理装置10は、クラス分類の結果得られたクラスコードを用いて予測係数メモリ15から予測係数を読み出す。この予測係数は予め学習によりクラス毎に対応して格納されており、オーディオ信号処理装置10はクラスコードに対応した予測係数を読み出すことにより、このときのオーディオ波形の特徴に合致した予測係数を用いることができる。

【0034】予測係数メモリ15から読み出された予測係数は、ステップSP14において予測演算部16の予測演算に用いられる。これにより、入力オーディオデータD10はその極性に応じた予測演算により、所望とするオーディオデータD16に変換される。かくして入力オーディオデータD10はその音質が改善されたオーディオデータD16に変換され、オーディオ信号処理装置10はステップSP15に移って当該処理手順を終了する。

【0035】次に、図1について上述した予測係数メモリ15に記憶するクラス毎の予測係数のセットを予め学習によって得るための学習回路について説明する。

【0036】図5において、学習回路30は、高音質の教師オーディオデータD30を生徒信号生成フィルタ37に受ける。生徒信号生成フィルタ37は、間引き率設定信号D39により設定された間引き率で教師オーディオデータD30を所定時間ごとに所定サンプル間引きくようになっている。

【0037】この場合、生徒信号生成フィルタ37における間引き率によって、生成される予測係数が異なり、これに応じて上述のオーディオ信号処理装置10で再現されるオーディオデータも異なる。例えば、上述のオーディオ信号処理装置10においてサンプリング周波数を高くすることでオーディオデータの音質を向上しようとする場合、生徒信号生成フィルタ37ではサンプリング周波数を減らす間引き処理を行う。また、これに対して

上述のオーディオ信号処理装置10において入力オーディオデータD10の欠落したデータサンプルを補うことで音質の向上を図る場合には、これに応じて、生徒信号生成フィルタ37ではデータサンプルを欠落させる間引き処理を行うようになされている。

【0038】かくして、生徒信号生成フィルタ37は教師オーディオデータ30から所定の間引き処理により生徒オーディオデータD37を生成し、これを極性判別部31、クラス分類部抽出部32及び予測演算部抽出部33にそれぞれ供給する。

【0039】極性判別部31は生徒信号生成フィルタ37から供給された生徒オーディオデータD37を所定時間毎の領域（この実施の形態の場合、例えば6サンプル毎とする）に分割した後、当該分割された各時間領域の波形について、その極性クラスを図2について上述したように分類する。

【0040】そして極性判別部31は生徒オーディオデータD37のこのとき分割された時間領域の極性判別結果を生徒オーディオデータD37の極性クラスデータD31としてクラス分類部34に供給する。

【0041】また、クラス分類部抽出部32は生徒信号生成フィルタ37から供給された生徒オーディオデータD37を、極性判別部31の場合と同様の時間領域（この実施の形態の場合例えば6サンプル）に分割することによりクラス分類しようとするオーディオ波形データD32を抽出し、これをクラス分類部34に供給する。

【0042】クラス分類部34は、クラス分類抽出部32において切り出されたオーディオ波形データD32について、当該オーディオ波形データD32を圧縮して圧縮データパターンを生成するADRC(Adaptive Dynamic Range Coding) 回路部と、オーディオ波形データD32の属するクラスコードを発生するクラスコード発生回路部とを有する。

【0043】ADRC回路部はオーディオ波形データD32に対して、例えば8ビットから2ビットに圧縮するような演算を行うことによりパターン圧縮データを形成する。このADRC回路部は、適応的量子化を行うものであり、ここでは、信号レベルの局所的なパターンを短い語長で効率的に表現することができるので、信号パターンのクラス分類のコード発生用に用いられる。

【0044】具体的には、オーディオ波形上の6つの8ビットのデータ（オーディオ波形データ）をクラス分類しようとする場合、 $2^8$ という膨大な数のクラスに分類しなければならず、回路上の負担が多くなる。そこで、この実施の形態のクラス分類部34ではその内部に設けられたADRC回路部で生成されるパターン圧縮データに基づいてクラス分類を行う。例えば6つのオーディオ波形データに対して1ビットの量子化を実行すると、6つのオーディオ波形データを6ビットで表すことができ、 $2^6 = 64$ クラスに分類することができる。

【0045】ここで、ADRC回路部は、切り出された領域内のオーディオ波形のダイナミックレンジをDR、ビット割り当をm、各オーディオ波形データのデータレベルをL、量子化コードをQとして、上述の(1)式と同様の演算により、領域内の最大値MAXと最小値MINとの間を指定されたビット長で均等に分割して量子化を行う。かくしてオーディオ波形上の6つの波形データが、それぞれ例えば8ビット（ $m=8$ ）で構成されるとすると、これらはADRC回路部においてそれぞれが2ビットに圧縮される。

【0046】このようにしてオーディオ波形のダイナミックレンジで正規化し圧縮されたオーディオ波形データをそれぞれ $q_n$ （ $n=1\sim6$ ）とすると、クラス分類部34に設けられたクラスコード発生回路部は、圧縮されたオーディオ波形データ $q_n$ に基づいて、上述の(2)式と同様の演算を実行することにより、そのブロック（ $q_1\sim q_n$ ）が属するクラスを示すクラスコードclassを算出し、当該算出されたクラスコードclassと極性判別部31により算出された極性クラス（CLASS0、CLASS1、CLASS2又はCLASS3）とを統合した後、当該統合されてなるクラスコードclass'を表すクラスコードデータD34を予測係数算出部36に供給する。因みに(2)式において、nは圧縮されたオーディオ波形データ $q_n$ の数を表し、この実施の形態の場合 $n=6$ であり、またPはビット割り当てを表し、この実施の形態の場合 $P=2$ である。

【0047】このようにして、クラス分類部34はクラスコードデータD34を生成し、これを予測係数算出部36に供給する。また、予測係数算出部36には、クラスコードデータD34に対応した時間軸領域のオーディオ波形データD33（ $x_1, x_2, \dots, x_n$ ）が予測演算部抽出部33において切り出されて供給される。

【0048】予測係数算出部36は、クラス分類部34から供給されたクラスコードclass'と、各クラスコードclass毎に切り出されたオーディオ波形データD33と、入力端T<sub>1n</sub>から供給された高音質の教師オーディオデータD30とを用いて、正規方程式を立てる。

【0049】すなわち、生徒オーディオデータD37のnサンプルのレベルをそれぞれ $x_1, x_2, \dots, x_n$ として、それぞれにpビットのADRCを行った結果の量子化データを $q_1, \dots, q_n$ とする。このとき、この領域のクラスコードclass'を上述の(2)式のように定義する。そして、上述のように生徒オーディオデータD37のレベルをそれぞれ、 $x_1, x_2, \dots, x_n$ とし、高音質の教師オーディオデータD30のレベルをyとしたとき、クラスコード毎に、予測係数 $w_1, w_2, \dots, w_n$ によるnタップの線形推定式を設定する。これを次式、

【0050】

【数4】

11

$$y = w_1 x_1 + w_2 x_2 + \dots + w_n x_n$$

【0051】とする。学習前は、 $w_n$  が未定係数である。

\* ブル数がMの場合、上述の(4)式に従って、次式、

【0053】

【0052】学習回路30では、クラスコード毎に、複数のオーディオデータに対して学習を行う。データサン

【数5】

$$y_k = w_1 x_{k1} + w_2 x_{k2} + \dots + w_n x_{kn}$$

..... (5)

【0054】が設定される。但し  $k=1, 2, \dots, M$  である。

※式、

【0056】

【0055】 $M > n$  の場合、予測係数  $w_1, \dots, w_n$  は一意的に決まらないので、誤差ベクトル  $e$  の要素を次

【数6】

※10

$$e_k = y_k - [w_1 x_{k1} + w_2 x_{k2} + \dots + w_n x_{kn}]$$

..... (6)

【0057】によって定義し(但し、 $k=1, 2, \dots, M$ )、次式、

★【0058】

★ 【数7】

$$e^2 = \sum_{k=0}^M e_k^2$$

..... (7)

【0059】を最小にする予測係数を求める。いわゆる、最小自乗法による解法である。

☆を求める。この場合、次式、

【0061】

【0060】ここで、(7)式による  $w_n$  の偏微分係数☆

【数8】

$$\frac{\partial e^2}{\partial w_i} = \sum_{k=0}^M 2 \left( \frac{\partial e_k}{\partial w_i} \right) e_k = \sum_{k=0}^M 2 x_{ki} \cdot e_k$$

$$= \sum_{k=0}^M 2 x_{ki} \cdot e_k \quad (i=1, 2, \dots, n)$$

..... (8)

【0062】を「0」にするように、各  $w_n$  ( $n=1 \sim 6$ ) を求めれば良い。

◆【0064】

【数9】

【0063】そして、次式、

◆

$$X_{ij} = \sum_{p=0}^M x_{pi} \cdot x_{pj}$$

..... (9)

【0065】

\* \* 【数10】

$$Y_i = \sum_{k=0}^M x_{ki} \cdot y_k$$

..... (10)

【0066】のように、 $X_{ij}$ 、 $Y_i$  を定義すると、(8)式は行列を用いて次式、

※【0067】

※ 【数11】

$$\begin{bmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & & \vdots \\ X_{m1} & X_{m2} & \dots & X_{mn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

..... (11)

【0068】として表される。

【0069】この方程式は、一般に正規方程式と呼ばれている。なお、ここでは  $n=6$  である。

【0070】全ての学習用データ(教師オーディオデータD30、クラスコードclass'、オーディオ波形データD33)の入力が完了した後、予測係数算出部36は各クラスコードclass'に上述の(11)式に示した正規方程式を立てて、この正規方程式を掃き出し法等の一

般的な行列解法を用いて、各  $w_n$  について解き、各クラスコード毎に、予測係数を算出する。予測係数算出部36は、算出された各予測係数(D36)を予測係数メモリ15に書き込む。

【0071】このような学習を行った結果、予測係数メモリ15には、量子化データ  $q_1, \dots, q_s$  で規定されるパターン毎に、高音質のオーディオデータ  $y$  を推定するための予測係数が、各クラスコード毎に格納され、

る。この予測係数メモリ15は、図1について上述したオーディオ信号処理装置10において用いられる。かかる処理により、線形推定式に従って通常のオーディオデータから高音質のオーディオデータを作成するための予測係数の学習が終了する。

【0072】このように、学習回路30は、オーディオ信号処理装置10において補間処理を行う程度を考慮して、生徒信号生成フィルタ37で高音質の教師オーディオデータの間引き処理を行うことにより、オーディオ信号処理装置10における補間処理のための予測係数を生成することができる。

【0073】以上の構成において、オーディオ信号処理装置10は、クラス分類部14のADRC処理においてオーディオ波形をそのダイナミックレンジで正規化することでオーディオ波形そのもののクラスコード classを得る。この場合、オーディオ波形のゼロレベル近傍及び大振幅部では音素が異なっている場合が多く、単にダイナミックレンジで正規化した結果でクラス分類を行うと、元々異なる音素であっても同一クラスと見なされてしまうことがある。従って、クラス分類部14では、オーディオ波形そのもののクラスコード classに、オーディオ波形の極性クラス CLASSを統合してクラスコード class' を算出し、これをクラス分類結果として予測演算に用いることにより、オーディオ波形そのものから得られたクラスコード classが同一クラスとなった場合でも、オーディオ波形の極性クラスに応じて確実にクラス分類することができる。

【0074】例えば、極性クラスが CLASS0又は CLASS3である場合、このことは切り出されたオーディオ波形データの値が全て正又は負であること、すなわち比較的大振幅の波形部分であることを表しており、また、極性クラスが CLASS1又は CLASS2である場合、このことは切り出されたオーディオ波形がゼロクロス部と正又は負を含む波形であること、すなわち比較的ゼロレベル近傍の波形部分であることを表しており、クラス分類部14はかかる極性クラスをオーディオ波形データそのもののクラスコード classに統合してクラス分類を行うことにより、異なる音素を異なるクラスコードとして分類することができる。

【0075】以上の構成によれば、入力オーディオデータD10の極性クラスを用いて入力オーディオデータD10をクラス分類し、当該クラス分類された結果に基づく予測係数を用いて予測演算するようにしたことにより、入力オーディオデータD10を一段と高音質のオーディオデータD16に変換することができる。

【0076】なお上述の実施の形態においては、オーディオ信号処理装置10及び学習装置30において、クラス分類部抽出部12、32及び予測演算部抽出部13、33により入力オーディオデータD10、D37を常に一定の範囲毎に切り出す場合について述べたが、本発明

はこれに限らず、例えば図1及び図5との対応部分に同一符号を付して示す図6及び図7に示すように、極性判別部11、31において算出された極性クラスに基づいて抽出制御信号CONT11、CONT31を可変クラス分類部抽出部12'、可変予測演算部抽出部13'及び可変クラス分類部抽出部32'、可変予測演算部抽出部33'に供給することにより入力オーディオデータD10、D37の切り出し範囲(タップ)を制御するようにしても良い。

【0077】この場合、極性判別部11、31は、極性クラス CLASS0、CLASS1、CLASS2及び CLASS3の頻度に基づいて切り出し範囲(タップの切り出し長)を制御することにより、タップの切り出し長を長くし過ぎることによる正極性のみ(CLASS0)又は負極性のみ(CLASS3)への分類頻度の低下を防止することができる。

【0078】この場合、オーディオデータの変換処理手順は図4との対応部分に同一符号を付して示す図8に示すように、オーディオ波形の極性を判別するステップSP12の次に、当該判別された極性に基づいて可変クラス分類部抽出部12'、32'及び可変予測演算部抽出部13'、33'におけるタップ抽出領域を制御する処理ステップSP21を挿入するようにすれば良い。

【0079】また上述の実施の形態においては、極性クラスとして4つの極性クラス CLASS0、CLASS1、CLASS2及び CLASS3を設ける場合について述べたが、本発明はこれに限らず、全て正の領域、全て負の領域、ゼロクロスを含む領域の3つの極性クラスに分類するようにしても良い。

【0080】また上述の実施の形態においては、予測方式として線形一次による手法を用いる場合について述べたが、本発明はこれに限らず、要は学習した結果を用いるようにすれば良く、例えば多次関数による手法等の種々の予測方式を適用することができる。

【0081】また上述の実施の形態においては、クラス分類部14においてADRCにより圧縮データパターンを生成する場合について述べたが、本発明はこれに限らず、可逆符号化(DPCM:Differential Pulse Code Modulation)又はベクトル量子化(VQ:Vector Quantize)等の圧縮手段を用いるようにしても良い。

【0082】また上述の実施の形態においては、学習回路30の生徒信号生成フィルタ37において教師オーディオデータD30から所定サンプル数を間引く場合について述べたが、本発明はこれに限らず、例えばビット数を削減する等、他の種々の方法を適用することができる。

【0083】

【発明の効果】上述のように本発明によれば、デジタルオーディオ信号の極性に基づいてデジタルオーディオ信号のクラスを分類し、当該分類されたクラスに対応した予測方式でデジタルオーディオ信号を変換するよ





【図3】

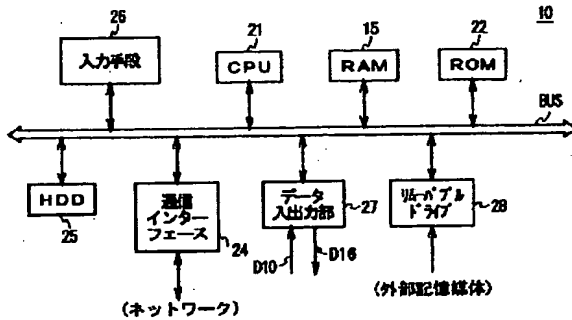


図3 オーディオ信号処理装置

【図5】

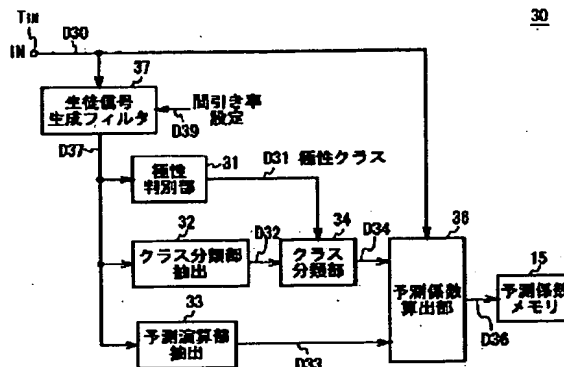


図5 学習回路の構成

【図6】

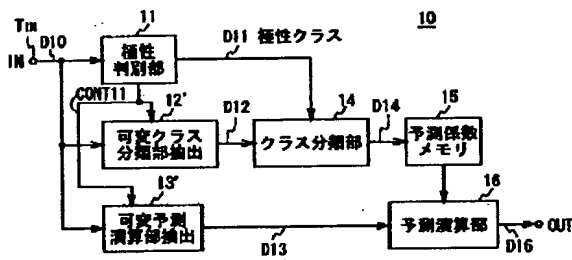


図6 他の実施の形態

【図7】

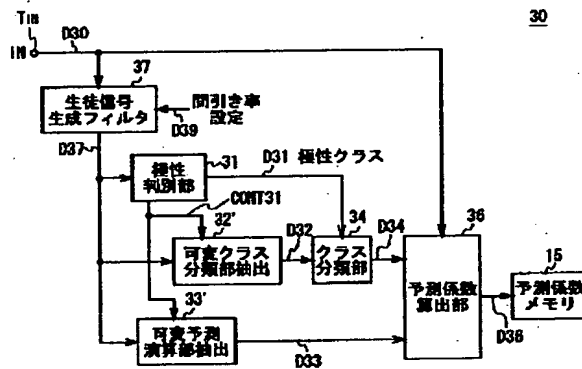


図7 他の実施の形態

フロントページの続き

(72)発明者 服部 正明  
東京都品川区北品川6丁目7番35号ソニー  
株式会社内

(72)発明者 木村 裕人  
東京都品川区北品川6丁目7番35号ソニー  
株式会社内

Fターム(参考) 5J064 AA01 BA04 BA06 BA13 BB03  
BB04 BC01 BC14 BD03

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-049383

(43)Date of publication of application : 15.02.2002

(51)Int.Cl.

G10L 13/00  
H03M 7/30

(21)Application number : 2000-238892

(71)Applicant : SONY CORP

(22)Date of filing : 02.08.2000

(72)Inventor : KONDO TETSUJIRO  
WATANABE TSUTOMU  
HATTORI MASAOKI  
KIMURA HIROTO

(54) DIGITAL SIGNAL PROCESSING METHOD AND LEARNING METHOD AND THEIR DEVICES, AND PROGRAM STORAGE MEDIUM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a digital signal processing method which further improves the waveform reproducibility of digital signals, a learning method and their devices, and a program storage medium.

SOLUTION: The classes of digital audio signals D10 are classified based on the polarities of the signals D10. Then, the signals D10 are converted by a predicting system corresponding to the class being classified. Thus, conversion that is made further suitable to the characteristics of the signals D10 can be performed.

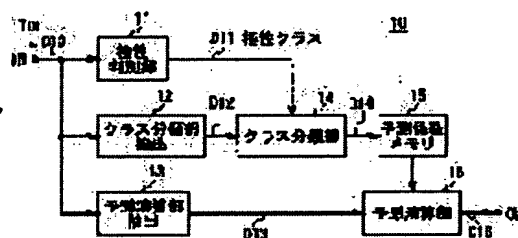


図1 オーディオ信号処理装置の構成

## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's  
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

**\* NOTICES \***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

**CLAIMS**

---

**[Claim(s)]**

[Claim 1] The digital-signal-processing equipment carry out having a prediction operation means generate a new digital audio signal [ come / by carrying out the prediction operation of the above-mentioned digital audio signal in the digital-signal-processing equipment which changes a digital audio signal by polar distinction means distinguish the polarity of the above-mentioned digital audio signal, class classification means classify the class of the above-mentioned digital audio signal based on the above-mentioned polar distinction result, and the prediction method corresponding to the class by which the classification was carried out / above-mentioned / in the above-mentioned digital audio signal / to change ] as the description.

[Claim 2] The above-mentioned polar distinction means is digital-signal-processing equipment according to claim 1 characterized by dividing the above-mentioned digital audio signal into a time-axis field, and distinguishing the polarity for every division field.

[Claim 3] The above-mentioned polar distinction means is digital-signal-processing equipment according to claim 1 characterized by carrying out a polar class division to three fields of the field where only a positive region contains the above-mentioned digital audio signal, and, as for it, only a negative region contains a zero cross at least.

[Claim 4] The above-mentioned prediction operation means is digital-signal-processing equipment according to claim 1 characterized by using the prediction coefficient currently generated by study based on the digital audio signal beforehand considered as a request.

[Claim 5] The digital-signal-processing approach of carrying out having the prediction operation step generate the polar distinction step which distinguishes the polarity of the above-mentioned digital audio signal, the class classification step into which the class of the above-mentioned digital audio signal classifies based on the above-mentioned polar distinction result, and a new digital audio signal [ come / by carrying out the prediction operation of the above-mentioned digital audio signal by the prediction method corresponding to the class by which the classification was carried out / above-mentioned / in the above-mentioned digital audio signal / to change ] in the digital-signal-processing approach of changing a digital audio signal as the description.

[Claim 6] At the above-mentioned polar distinction step, the above-mentioned digital audio signal is the digital-signal-processing approach according to claim 5 characterized by being divided into a time-axis field and distinguishing the polarity for every division field.

[Claim 7] Only for a positive region, at the above-mentioned polar distinction step, only a negative region is [ the above-mentioned digital audio signal ] the digital-signal-processing approach according to claim 5 that it is characterized by carrying out a polar class division to three fields of the field containing a zero cross, at least.

[Claim 8] The digital-signal-processing approach according to claim 5 characterized by using the prediction coefficient currently generated by study at the above-mentioned prediction operation step based on the digital audio signal beforehand considered as a request.

[Claim 9] In the study equipment which generates the prediction coefficient used for the prediction operation of the above-mentioned transform processing of the digital-signal-processing equipment which changes a digital audio signal A student digital audio signal

generation means to generate the student digital audio signal which degraded the digital audio signal concerned from the digital audio signal considered as a request, A polar distinction means to distinguish the polarity of the above-mentioned student digital audio signal, A class classification means to classify the class of the above-mentioned student digital audio signal based on the polarity by which distinction was carried out [ above-mentioned ], Study equipment characterized by having a prediction coefficient calculation means to compute the prediction coefficient corresponding to the above-mentioned class based on the above-mentioned digital audio signal and the above-mentioned student digital audio signal.

[Claim 10] The above-mentioned polar distinction means is study equipment according to claim 9 characterized by dividing the above-mentioned digital audio signal into a time-axis field, and distinguishing the polarity for every division field.

[Claim 11] The above-mentioned polar distinction means is study equipment according to claim 9 characterized by carrying out a polar class division to three fields of the field where only a positive region contains the above-mentioned digital audio signal, and, as for it, only a negative region contains a zero cross at least.

[Claim 12] In the study approach which generates the prediction coefficient used for the prediction operation of the above-mentioned transform processing of the digital-signal-processing equipment which changes a digital audio signal The student digital audio signal generation step which generates the student digital audio signal which degraded the digital audio signal concerned from the digital audio signal considered as a request, The polar distinction step which distinguishes the polarity of the above-mentioned student digital audio signal, The class classification step into which the class of the above-mentioned student digital audio signal is classified based on the polarity by which distinction was carried out [ above-mentioned ], The study approach characterized by having the prediction coefficient calculation step which computes the prediction coefficient corresponding to the above-mentioned class based on the above-mentioned digital audio signal and the above-mentioned student digital audio signal.

[Claim 13] At the above-mentioned polar distinction step, the above-mentioned digital audio signal is the study approach according to claim 12 characterized by being divided into a time-axis field and distinguishing the polarity for every division field.

[Claim 14] Only for a positive region, at the above-mentioned polar distinction step, only a negative region is [ the above-mentioned digital audio signal ] the study approach according to claim 12 that it is characterized by carrying out a polar class division to three fields of the field containing a zero cross, at least.

[Claim 15] The program storing medium which makes digital-signal-processing equipment perform the program contain the prediction operation step which generates in the new digital audio signal which comes to change the above-mentioned digital audio signal by carrying out the prediction operation of the above-mentioned digital audio signal using the polar distinction step which distinguishes the polarity of a digital audio signal, the class classification step into which the class of the above-mentioned digital audio signal classifies based on the above-mentioned polar distinction result, and the prediction coefficient corresponding to the class by which the classification was carried out [ above-mentioned ].

[Claim 16] The student digital audio signal generation step which generates the student digital audio signal which degraded the digital audio signal concerned from the digital audio signal considered as a request, The polar distinction step which distinguishes the polarity of the above-mentioned student digital audio signal, The class classification step into which the class of the above-mentioned student digital audio signal is classified based on the polarity by which distinction was carried out [ above-mentioned ], The program storing medium which makes study equipment perform the program containing the prediction coefficient calculation step which computes the prediction coefficient corresponding to the above-mentioned class based on the above-mentioned digital audio signal and the above-mentioned student digital audio signal.

---

[Translation done.]

\* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to a program storing medium at the digital-signal-processing approach, the study approaches, and those equipment lists, and is a rate converter or PCM (Pulse Code Modulation). It applies to a program storing medium and is suitable for the digital-signal-processing approach of performing interpolation processing of data to a digital signal in decode equipment etc., the study approaches, and those equipment lists.

[0002]

[Description of the Prior Art] Before inputting a digital audio signal into digital one/analog converter conventionally, exaggerated sampling processing which changes a sampling frequency by several times the original value is performed. Thereby, the digital audio signal outputted from digital one/analog converter is made as [ eliminate / the phase characteristic of an analog anti-aliasing filter is kept constant in an audio frequency quantity region, and / the effect of the image noise of the digital system accompanying a sampling ].

[0003] In this exaggerated sampling processing, the digital filter of a linearity primary (straight line) interpolation method is usually used. Such a digital filter generates linear interpolation data in quest of the average value of two or more existing data, when a sampling rate changes or data are missing.

[0004]

[Problem(s) to be Solved by the Invention] However, although the amount of data has become precise several times to time amount shaft orientations by linearity linear interpolation as for the digital audio signal after exaggerated sampling processing, the frequency band of the digital audio signal after exaggerated sampling processing seldom changes to before conversion, and the tone quality itself has not improved. Furthermore, since the interpolated data were not necessarily generated based on the wave of the analog audio signal in front of A/D conversion, most its wave repeatability has not improved.

[0005] Moreover, although the frequency was changed using the sampling rate converter when the digital audio signal from which a sampling frequency differs was dubbed, it was difficult to be able to perform only interpolation of linear data but to improve tone quality and wave repeatability with a linearity primary digital filter, also by this case. Furthermore, it is the same when the data sample of a digital audio signal is missing.

[0006] This invention was made in consideration of the above point, and tends to propose a program storing medium in the digital-signal-processing approach which may improve the wave repeatability of a digital signal much more, the study approaches, and those equipment lists.

[0007]

[Means for Solving the Problem] Since this technical problem is solved, conversion which was adapted for the description of a digital audio signal much more can be performed by classifying the class of a digital audio signal in this invention based on the polarity of a digital audio signal, and having changed the digital audio signal by the prediction method corresponding to the classified class concerned.

[0008]

[Embodiment of the Invention] About a drawing, the gestalt of 1 operation of this invention is explained in full detail below.

[0009] In case the audio signal processor 10 raises the sampling rate of a digital audio signal (this is called audio data below) or audio data are interpolated in drawing 1, it is made as [ generate / the audio data near a true value / class classification application processing ]. Incidentally, a digital audio signal means the sound signal showing the voice which people and an animal utter, the musical-sound signal showing the musical sound which a musical instrument emits, and the signal showing other sounds.

[0010] That is, in the audio signal processor 10, the polar distinction section 11 distinguishes that polar class about the wave of each divided time domain concerned by the polar distinction approach shown in drawing 2, after dividing into the field (it carries out to every the case of the gestalt of this operation, 6 [ for example, ], samples) for every predetermined time the input audio data D10 shown in drawing 2 supplied from the input terminal TIN.

[0011] Namely, when all the taps started like the divided field AR 1 (current data C1) in drawing 2 are forward, This polar class While a zero cross exists in the field which set to CLASS0 and was started like the divided field AR 2 (current data C2), when current data (C2) are forward, This polar class While a zero cross exists in the started field which set to CLASS1 and was started like the divided field AR 4 (current data C4), when current data (C4) are negative, This polar class When all the taps that set to CLASS2 and were started like the divided field AR 3 (current data C3) are negative, it is this polar class. It is referred to as CLASS3.

[0012] Thus, a phoneme in case the audio data D10 are near the zero level, and the phoneme in the large amplitude section are clearly distinguishable much more in the field of both positive/negative by setting up the polar class based on the polarity of the audio data D10.

[0013] The polar distinction section 11 is supplied to the class classification section 14 by using as the polar class data D11 the polar distinction result (CLASS0, CLASS1, CLASS2, or CLASS3) searched for corresponding to the current data at this time of the input audio data D10.

[0014] Moreover, by dividing into the same time domain as the case of polar distinction \*\*\*\* 11, the cases, for example, six samples, of the gestalt of this operation, the input audio data D10 supplied from the input terminal TIN, the class classification section extract section 12 extracts the audio data point D12 which is going to carry out a class classification, and supplies this to the class classification section 14.

[0015] The class classification section 14 is ADRC (Adaptive Dynamic Range Coding) which compresses the audio data point D12 concerned, and generates a compression data pattern about the audio data point D12 started in the class classification extract section 12. It has the circuit section and the class code generating circuit section which generates the class code to which the audio data point D12 belongs.

[0016] The ADRC circuit section forms pattern compressed data by performing an operation which is compressed into 2 bits from 8 bits as opposed to the audio data point D12. Since this ADRC circuit section can perform accommodative quantization and can express the local pattern of signal level efficiently by the short word length here, it is used for code generating of a class classification of a signal pattern.

[0017] When it is going to carry out the class classification of the six 8-bit data on an audio wave (audio data point), it must classify into a huge number 248 of classes, and, specifically, the burden on a circuit increases. So, in the class classification section 14 of the gestalt of this operation, a class classification is performed based on the pattern compressed data generated in the ADRC circuit section prepared in that interior. For example, if 1-bit quantization is performed to six audio data points, six audio data points can be expressed with 6 bits, and it can classify into  $2^6 = 64$  class.

[0018] Here, the ADRC circuit section is a degree type and [0019], when the data level of m and each audio data point is set to L and a quantization code is set [ the dynamic range of an audio wave in the started field ] to Q for this [ DR and / bit rate ].

[Equation 1]



$$DR = MAX - MIN + 1$$

$$Q = \{ (L - MIN + 0.5) \times 2^m / DR \} \dots\dots (1)$$

[0020] It is alike, and it follows and quantizes by dividing equally by the bit length which had between the maximum MAX in a field, and the minimum values MIN specified. In addition, in (1) type, { } means the cut-off processing below decimal point. In this way, supposing six data points on an audio wave consist of 8 bits ( $m=8$ ), for example, respectively, as for these, each will be compressed into 2 bits in the ADRC circuit section.

[0021] Thus, if the audio data point normalized and compressed with the dynamic range is set to  $q_n$  ( $n=1-6$ ), respectively, the class code generating circuit section prepared in the class classification section 14 is the compressed audio data point  $q_n$ . It is based and is a degree type and [0022].

[Equation 2]

$$class = \sum_{i=1}^n q_i (2^P)^i \dots\dots (2)$$

[0023] Class code class which shows the class to which the block ( $q_1 - q_6$ ) belongs by being alike and performing the shown operation Class code based on the computed audio data point D12 concerned while computing It is an above-mentioned polar class to class. The class code concerned integrated after unifying CLASS The class code data D14 showing class' are supplied to the prediction coefficient memory 15. This class code class ' shows the read-out address at the time of reading a prediction coefficient from the prediction coefficient memory 15.

Incidentally it is the audio data point  $q_n$  into which  $n$  was compressed in (2) types. A number is expressed and, in the case of the gestalt of this operation, it is  $n=6$ , and  $P$  expresses bit assignment and, in the case of the gestalt of this operation, is  $P=2$ .

[0024] Thus, the class classification section 14 is the class code of audio data-point D12 themselves started from the input audio data D10 in the class classification section extract section 12. class and polar class of the audio data point D12 The class code data (class') D14 which unified CLASS are generated, and this is supplied to the prediction coefficient memory 15. Incidentally, it is the polar class of the class code class of audio data-point D12 themselves, and the audio data point D12. As an approach of unifying CLASS, the class classification section 14 is the class code of for example, audio data-point D12 themselves. It is a polar class to class. These can be unified by adding CLASS.

[0025] Set  $w_1 - w_n$  of the prediction coefficient which the set of the prediction coefficient corresponding to each class code is memorized to the address corresponding to a class code by the prediction coefficient memory 15, respectively, and is memorized to the address corresponding to the class code concerned based on the class code data D14 supplied from the class classification section 14 It is read and the prediction operation part 16 is supplied.

[0026] The prediction operation part 16 is the audio data point (prediction tap) D13 ( $x_1 - x_n$ ) which was started in the time-axis field from the input audio data D10 in the prediction operation part extract section 13 and which is going to carry out a prediction operation, and prediction coefficient  $w_1 - w_n$ . It receives and is a degree type [0027].

[Equation 3]

$$y' = w_1 x_1 + w_2 x_2 + \dots\dots + w_n x_n \dots\dots (3)$$

[0028] Prediction result  $y'$  is obtained by being alike and performing the shown sum-of-products operation. This forecast  $y'$  is outputted from the prediction operation part 16 as audio data D16 with which tone quality has been improved.

[0029] In addition, although functional block mentioned above about drawing 1 as a configuration of the audio signal processor 10 was shown, the equipment of a computer configuration shown in drawing 3 in the gestalt of this operation as a concrete configuration which constitutes this functional block is used. In drawing 3 namely, the audio signal processor 10 It has RAM (Random Access Memory)15 which constitutes CPU21, ROM (Read Only Memory)22, and the prediction coefficient memory 15 through Bus BUS, and the configuration to which each circuit section was

connected, respectively. By performing the various programs stored in ROM22, CPU11 It is made as [ operate / as each functional block (the polar distinction section 11, the class classification section extract section 12, the prediction operation part extract section 13, the class classification section 14, and prediction operation part 16) mentioned above about drawing 1 ].

[0030] Moreover, it has the removable drive 28 which reads information from external storage, such as the communication link interface 24 and floppy disk which communicate between networks, and a magneto-optic disk, to the audio signal processor 10, each program for performing class classification application processing mentioned above about drawing 1 from a network course or external storage can be read to the hard disk of a hard disk drive unit 25, and class classification adaptation processing can also be performed according to \*\*\*\*\* and the read program concerned.

[0031] A user performs class classification processing mentioned above about drawing 1 to CPU21 by inputting various commands through the input means 26, such as a keyboard and a mouse. In this case, after the audio signal processor 10 inputs the audio data (input audio data) D10 which are going to raise tone quality through the data I/O section 27 and performs class classification application processing to the input audio data D10 concerned, it is made as [ output / through the data I/O section 27 / the audio data D16 whose tone quality improved / outside ].

[0032] Incidentally, drawing 4 shows the procedure of the class classification adaptation processing in the audio signal processor 10, and if the audio signal processor 10 goes into the procedure concerned from a step SP 11, it will compute the polarity of the input audio data D10 in the polar distinction section 11 in the continuing step SP 12.

[0033] This computed polarity is also that for ensuring the class classification of the OTIO data point D12 much more, and the audio signal processor 10 carries out the class classification of the audio data point D12 based on the audio data point D12 and the polar class D11 in a step SP 13 by the class classification section 14. And the audio signal processor 10 reads a prediction coefficient from the prediction coefficient memory 15 using the class code obtained as a result of the class classification. This prediction coefficient corresponds for every class by study beforehand, and is stored, and the audio signal processor 10 can use the prediction coefficient corresponding to the description of an audio wave at this time by reading the prediction coefficient corresponding to a class code.

[0034] The prediction coefficient read from the prediction coefficient memory 15 is used for the prediction operation of the prediction operation part 16 in a step SP 14. Thereby, the input audio data D10 are changed into the audio data D16 considered as a request by the prediction operation according to the polarity. In this way, the input audio data D10 are changed into the audio data D16 with which the tone quality has been improved, and the audio signal processor 10 moves to a step SP 15, and ends the procedure concerned.

[0035] Next, the study circuit for obtaining beforehand the set of the prediction coefficient for every class memorized to the prediction coefficient memory 15 mentioned above about drawing 1 by study is explained.

[0036] In drawing 5, the study circuit 30 receives the teacher audio data D30 of the quality of loud sound in the student signal generation filter 37. The student signal generation filter 37 is made as [ lengthen / the teacher audio data D30 / for every predetermined time / at the rate of infanticide set up by the rate setting signal D39 of infanticide / between predetermined samples ].

[0037] In this case, the prediction coefficient generated changes with rates of infanticide in the student signal generation filter 37, and the audio data reproduced with the above-mentioned audio signal processor 10 according to this also differ. For example, when it is going to improve the tone quality of audio data by making a sampling frequency high in the above-mentioned audio signal processor 10, infanticide processing which reduces a sampling frequency is performed with the student signal generation filter 37. Moreover, when aiming at improvement in tone quality by compensating the data sample which lacked the input audio data D10 in the above-mentioned audio signal processor 10 to this, according to this, it is made as [ perform / infanticide processing made to lack a data sample ] with the student signal generation filter 37.

[0038] In this way, the student signal generation filter 37 generates the student audio data D37 by predetermined infanticide processing from the teacher audio data 30, and supplies this to the polar distinction section 31, the class classification section extract section 32, and the prediction operation part extract section 33, respectively.

[0039] After the polar distinction section 31 divides into the field (it carries out to every the case of the gestalt of this operation, 6 [ for example, ], samples) for every predetermined time the student audio data D37 supplied from the student signal generation filter 37, as that polar class was mentioned above about drawing 2, it is classified about the wave of each divided time domain concerned.

[0040] And the polar distinction section 31 supplies the polar distinction result of the time domain divided at this time of the student audio data D37 to the class classification section 34 as polar class data D31 of the student audio data D37.

[0041] Moreover, by dividing into the same time domain as the case of the polar distinction section 31, the cases, for example, six samples, of the gestalt of this operation, the student audio data D37 supplied from the student signal generation filter 37, the class classification section extract section 32 extracts the audio data point D32 which is going to carry out a class classification, and supplies this to the class classification section 34.

[0042] The class classification section 34 is ADRC (Adaptive Dynamic Range Coding) which compresses the audio data point D32 concerned, and generates a compression data pattern about the audio data point D32 started in the class classification extract section 32. It has the circuit section and the class code generating circuit section which generates the class code to which the audio data point D32 belongs.

[0043] The ADRC circuit section forms pattern compressed data by performing an operation which is compressed into 2 bits from 8 bits as opposed to the audio data point D32. Since this ADRC circuit section can perform accommodative quantization and can express the local pattern of signal level efficiently by the short word length here, it is used for code generating of a class classification of a signal pattern.

[0044] When it is going to carry out the class classification of the six 8-bit data on an audio wave (audio data point), it must classify into a huge number 248 of classes, and, specifically, the burden on a circuit increases. So, in the class classification section 14 of the gestalt of this operation, a class classification is performed based on the pattern compressed data generated in the ADRC circuit section prepared in that interior. For example, if 1-bit quantization is performed to six audio data points, six audio data points can be expressed with 6 bits, and it can classify into  $2^6 = 64$  class.

[0045] Here, the ADRC circuit section sets [ the dynamic range of an audio wave in the started field ] a quantization code to Q for this [ DR and / bit rate ], setting the data level of m and each audio data point as L, and quantizes by dividing equally by the bit length which had between the maximum MAX in a field, and the minimum values MIN specified by the same operation as above-mentioned (1) type. In this way, supposing six data points on an audio wave consist of 8 bits ( $m = 8$ ), for example, respectively, as for these, each will be compressed into 2 bits in the ADRC circuit section.

[0046] Thus, if the audio data point normalized and compressed with the dynamic range of an audio wave is set to  $q_n$  ( $n=1-6$ ), respectively The class code generating circuit section prepared in the class classification section 34 Based on the compressed audio data point  $q_n$ , by performing the same operation as above-mentioned (2) types Class code class which shows the class to which the block ( $q_1 - q_6$ ) belongs It computes. The computed class code class concerned After unifying the polar class (CLASS0, CLASS1, CLASS2, or CLASS3) computed by the polar distinction section 31, The class code concerned which it comes to unify The class code data D34 showing class' are supplied to the prediction coefficient calculation section 36. Incidentally it is the audio data point  $q_n$  into which n was compressed in (2) types. A number is expressed and, in the case of the gestalt of this operation, it is  $n=6$ , and P expresses bit assignment and, in the case of the gestalt of this operation, is  $P=2$ .

[0047] Thus, the class classification section 34 generates the class code data D34, and supplies this to the prediction coefficient calculation section 36. Moreover, the audio data point D33 ( $x_1$ ,

$x_2, \dots, x_n$ ) of the time-axis field corresponding to the class code data D34 is started and supplied to the prediction coefficient calculation section 36 in the prediction operation part extract section 33.

[0048] class code class' to which the prediction coefficient calculation section 36 was supplied from the class classification section 34, and each class code class every — a normal equation is stood using the started audio data point D33 and the teacher audio data D30 of the quality of loud sound supplied from the input edge TIN.

[0049] namely, the level of  $n$  sample of the student audio data D37 — respectively —  $x_1, x_2, \dots, x_n$  \*\*\*\*\* — the quantization data of the result of having been alike, respectively and having performed  $p$ -bit ADRC —  $q_1, \dots, q_n$  \*\* — it carries out. At this time, class code class' of this field is defined like above-mentioned (2) types. And they are  $x_1, x_2, \dots, x_n$  about the level of the student audio data D37 as mentioned above, respectively. When it carries out and level of the teacher audio data D30 of the quality of loud sound is set to  $y$ , they are a prediction coefficient  $w_1$ , and  $w_2, \dots, w_n$  for every class code. The linearity presumption type of  $n$  tap to depend is set up. It is this A degree type and [0050]

[Equation 4]

$$y = w_1 x_1 + w_2 x_2 + \dots + w_n x_n \quad \dots (4)$$

[0051] It carries out. Before study, it is  $w_n$ . It is an undetermined coefficient.

[0052] In the study circuit 30, it learns to two or more audio data for every class code. When a data measurement size is  $M$ , above-mentioned (4) types are followed, and it is a degree type and [0053].

[Equation 5]

$$y_k = w_1 x_{k1} + w_2 x_{k2} + \dots + w_n x_{kn} \quad \dots (5)$$

[0054] It \*\*\*\*\* however,  $k = 1, 2, \dots$  it is  $M$ .

[0055] In  $M > n$ , they are a prediction coefficient  $w_1$  and  $\dots w_n$ . Since it is not decided uniquely, it is the element of the error vector  $e$  A degree type and [0056]

[Equation 6]

$$e_k = y_k - \{w_1 x_{k1} + w_2 x_{k2} + \dots + w_n x_{kn}\} \quad \dots (6)$$

[0057] It is defined as alike (however,  $k = 1, 2, \dots, M$ ), and is a degree type and [0058].

[Equation 7]

$$e^2 = \sum_{k=0}^M e_k^2 \quad \dots (7)$$

[0059] It asks for the prediction coefficient made into min. It is a solution method by the so-called least square method.

[0060]  $w_n$  according to (7) types here It asks for a partial differential coefficient. In this case, a degree type, [0061]

[Equation 8]

$$\begin{aligned} \frac{\partial e^2}{\partial w_i} &= \sum_{k=0}^M 2 \left( \frac{\partial e_k}{\partial w_i} \right) e_k = \sum_{k=0}^M 2 x_{ki} \cdot e_k \\ &= \sum_{k=0}^M 2 x_{ki} \cdot e_k \quad (i = 1, 2, \dots, n) \quad \dots (8) \end{aligned}$$

[0062] What is necessary is just to calculate each  $w_n$  ( $n=1-6$ ) so that it may be made "0."

[0063] And a degree type, [0064]

[Equation 9]

$$X_{ij} = \sum_{p=0}^M x_{pi} \cdot x_{pj} \quad \dots (9)$$

[0065]

[Equation 10]

$$Y_i = \sum_{k=0}^M X_{ki} \cdot y_k \quad \dots\dots (10)$$

[0066] \*\* -- like --  $X_{ij}$  and  $Y_i$  if a definition is given -- (8) types -- a matrix -- using -- a degree type and [0067]

[Equation 11]

$$\begin{bmatrix} X_{11} & X_{12} & \dots\dots\dots & X_{1n} \\ X_{21} & X_{22} & \dots\dots\dots & X_{2n} \\ \vdots & \vdots & & \vdots \\ X_{m1} & X_{m2} & \dots\dots\dots & X_{mn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad \dots\dots (11)$$

[0068] It is expressed by carrying out.

[0069] Generally this equation is called the normal equation. In addition, it is  $n=6$  here.

[0070] the normal equation having shown the prediction coefficient calculation section 36 at above-mentioned (11) ceremony at each class code class ' after the input of all the data for study (the teacher audio data D30, class code class ', audio data point D33) was completed -- standing -- this normal equation -- sweeping out -- general matrix solution methods, such as law, -- using -- every --  $W_n$  \*\*\*\*\* -- it solves and a prediction coefficient is computed for every class code. The prediction coefficient calculation section 36 writes each computed prediction coefficient (D36) in the prediction coefficient memory 15.

[0071] As a result of performing such study, in the prediction coefficient memory 15, they are the quantization data  $q_1, \dots, q_6$ . The prediction coefficient for presuming the audio data  $y$  of the quality of loud sound is stored for every class code for every pattern specified. This prediction coefficient memory 15 is used in the audio signal processor 10 mentioned above about drawing 1. By this processing, study of the prediction coefficient for creating the audio data of the quality of loud sound from the usual audio data according to a linearity presumption type is completed.

[0072] Thus, the study circuit 30 can generate the prediction coefficient for the interpolation processing in the audio signal processor 10 in consideration of extent which performs interpolation processing in the audio signal processor 10 by performing infanticide processing of the teacher audio data of the quality of loud sound with the student signal generation filter 37.

[0073] The audio signal processor 10 is the class code of the audio wave itself by normalizing [ in / on the above configuration and / ADRC processing of the class classification section 14 ] an audio wave with the dynamic range. class is obtained. In this case, in the large amplitude section near the zero level of an audio wave, if a class classification is performed by the result which phonemes differ in many cases and was only normalized with the dynamic range, even if it is a different phoneme from the first, it may be regarded as the same class. Therefore, at the class classification section 14, it is the class code of the audio wave itself. Polar class of an audio wave [ class ] CLASS is unified and it is a class code. Class code obtained from the audio wave itself by computing class' and using for a prediction operation by making this into a class classification result Even when class becomes the same class, according to the polar class of an audio wave, a class classification can be carried out certainly.

[0074] for example, polar class CLASS0 -- or -- when it is CLASS3, all the values of the audio data point by which this was started are forward or negative -- namely, it is the wave part of the large amplitude comparatively -- expressing -- \*\*\*\* -- moreover, polar class CLASS1 -- or -- when it is CLASS2, this is a wave in which the started audio wave contains the zero cross section, forward, or negative -- That is, it expresses that it is a wave part near the zero level comparatively, and the class classification section 14 is the class code of the audio data point

itself about this polar class. By unifying to class and performing a class classification A different phoneme can be classified as a different class code.

[0075] According to the above configuration, the input audio data D10 are convertible for the audio data D16 of the quality of loud sound much more by carrying out the class classification of the input audio data D10 using the polar class of the input audio data D10, and having been made to carry out a prediction operation using the prediction coefficient based on the result concerned by which the class classification was carried out.

[0076] In addition, although the case where the input audio data D10 and D37 were always cut down for every fixed range by the class classification section extract sections 12 and 32 and the prediction operation part extract sections 13 and 33 was described in the audio signal processor 10 and study equipment 30 in the gestalt of above-mentioned operation As this invention is shown in drawing 6 and drawing 7 which attach and show the same sign to a corresponding point not only with this but drawing 1 and drawing 5 It is based on the polar class computed in the polar distinction sections 11 and 31. The extract control signals CONT11 and CONT31 Adjustable class classification section extract section 12', You may make it control the logging range of the input audio data D10 and D37 (tap) by supplying adjustable prediction operation part extract section 13' and adjustable class classification section extract section 32', and adjustable prediction operation part extract section 33'.

[0077] in this case, the polar distinction sections 11 and 31 -- polar class CLASS0, CLASS1, and CLASS2 -- and -- Only as for the straight polarity by lengthening the logging length of a tap too much, by starting based on the frequency of CLASS3 and controlling the range (logging length of a tap), only negative polarity can prevent the fall of the classification frequency to (CLASS3) (CLASS0).

[0078] In this case, what is necessary is just to make it insert the processing step SP 21 which controls the tap extract field in adjustable class classification section extract section 12', 32' and adjustable prediction operation part extract section 13', and 33' based on the distinguished polarity concerned, next the transform-processing procedure of audio data is the step SP 12 which distinguishes the polarity of an audio wave, as shown in drawing 8 which attaches and shows the same sign to a corresponding point with drawing 4.

[0079] moreover, the gestalt of above-mentioned operation -- setting -- as a polar class -- four polar classes CLASS0, CLASS1, and CLASS2 -- and -- Although the case where CLASS3 was formed was described, you may make it classify this invention into three polar classes of a forward field, and a negative field and the field containing a zero cross altogether not only in this but altogether.

[0080] moreover, the result of in short having learned this invention not only in this although the case where the technique by linearity primary was used as a prediction method was described in the gestalt of above-mentioned operation -- using -- it makes -- \*\*\*\*\* -- for example, -- many -- various prediction methods, such as technique by degree function, are applicable.

[0081] moreover -- although the case where ADRC generated a compression data pattern in the class classification section 14 was described in the gestalt of above-mentioned operation -- this invention -- not only this but reversible coding (DPCM:Differential Pulse Code Modulation) Or vector quantization (VQ:Vector Quantize) etc. -- you may make it use a compression means

[0082] Moreover, in the gestalt of above-mentioned operation, although the case where a predetermined measurement size was thinned out from the teacher audio data D30 in the student signal generation filter 37 of the study circuit 30 was described, this invention can apply other various approaches, such as reducing not only this but the number of bits.

[0083]

[Effect of the Invention] According to this invention, conversion which was adapted for the description of a digital audio signal much more can be performed as mentioned above by classifying the class of a digital audio signal based on the polarity of a digital audio signal, and having changed the digital audio signal by the prediction method corresponding to the classified class concerned.

[Translation done.]

**\* NOTICES \***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

**DESCRIPTION OF DRAWINGS**

---

**[Brief Description of the Drawings]**

**[Drawing 1]** It is the block diagram showing the configuration of the digital-signal-processing equipment by this invention.

**[Drawing 2]** It is the signal waveform diagram with which explanation of polar distinction is presented.

**[Drawing 3]** It is the block diagram showing the configuration of an audio signal processor.

**[Drawing 4]** It is the flow chart which shows an audio signal transform-processing procedure.

**[Drawing 5]** It is the block diagram showing the configuration of the study equipment by this invention.

**[Drawing 6]** It is the block diagram showing the gestalt of other operations of digital-signal-processing equipment.

**[Drawing 7]** It is the block diagram showing the gestalt of other operations of study equipment.

**[Drawing 8]** It is the flow chart which shows the audio signal transform-processing procedure by the gestalt of other operations.

**[Description of Notations]**

10 [ .. A prediction coefficient memory, 16 / .. Prediction operation part, 36 / .. The prediction coefficient calculation section, 37 / .. Student signal generation filter. ] .... 11 An audio signal processor, 31 .. 14 The polar distinction section, 34 .. The class classification section, 15

---

**[Translation done.]**